JOURNAL OF MULTIVARIATE ANALYSIS 18, 1–31 (1986)

# Third Order Asymptotic Properties of Maximum Likelihood Estimators for Gaussian ARMA Processes*

MASANOBU TANIGUCHI

*Hiroshima University, Hiroshima, Japan*

*Communicated by E. J. Hannan*

In this paper we investigate various third-order asymptotic properties of maximum likelihood estimators for Gaussian ARMA processes by the third-order Edgeworth expansions of the sampling distributions. We define a third-order asymptotic efficiency by the highest probability concentration around the true value with respect to the third-order Edgeworth expansion. Then we show that the maximum likelihood estimator is not always third-order asymptotically efficient in the class $A_3$ of third-order asymptotically median unbiased estimators. But, if we confine our discussions to an appropriate class $D$ ($\subset A_3$) of estimators, we can show that appropriately modified maximum likelihood estimator is always third-order asymptotically efficient in $D$. 

## 1. INTRODUCTION

Recently some systematic studies of higher order asymptotic efficiency for stationary processes have been developed. For an AR(1) process, Akahira [1] showed that appropriately modified least squares estimator of the first-order coefficient $\theta$ is second-order asymptotically efficient in the class $A_2$ of second order asymptotically median unbiased (AMU) estimators if efficiency is measured by the degree of concentration of the sampling distribution up to second order. This concept of efficiency was introduced by Akahira [1] and Akahira and Takeuchi [3], and these results are briefly reviewed in Section 2. Also Akahira [2] showed the second-order asymptotic efficiency of a modified maximum likelihood estimator (MLE) of $\theta$ for AR(1) case.

Received March 15, 1984; revised February 20, 1985.

AMS 1980 subject classifications. Primary 62F12, 62M15: secondary 62M10, 62E20.

Key words and phrases: Gaussian autoregressive moving average processes, spectral density, Toeplitz matrix, maximum likelihood estimator, third order asymptotic efficiency, Edgeworth expansion, residue theorem

* This work was partially supported by Japan Society of Promotion of Science, United States-Japan Cooperative Science Program and Grant-in-Aid of the Ministry of Japan.



0047-259X/86 $3.00

Now, let $\{X_t;\ t=0, \pm 1, \pm 2,...,\}$ be a Gaussian ARMA process with spectral density $f_\theta(\lambda)$ which depends on an unknown parameter $\theta$. In this case Taniguchi [15] showed that appropriately modified MLE of $\theta$ is second-order asymptotically efficient in the class $A_2$. In this paper, furthermore we shall develop the results of Taniguchi [15] for third-order case, and investigate various third-order asymptotic properties of the MLE for Gaussian ARMA processes. Our results grasp many other works done for AR(1) or MA(1) processes as special cases.

In Section 5, we shall derive the third order-bound distribution for the class $A_3$ of third order AMU estimators. This bound distribution gives the highest probability concentration around the true value among estimators in $A_3$. If an estimator in $A_3$ attains the third-order bound distribution we call it third order asymptotically efficient in $A_3$.

In Section 6, we show that the maximum likelihood estimators for ARMA process are not always third order asymptotically efficient in $A_3$. While we give the necessary and sufficient condition for the spectral density such that appropriately modified maximum likelihood estimator is third order asymptotically efficient in $A_3$. Also we give the Edgeworth expansions and the mean square errors of the maximum likelihood estimators up to third order.

In Section 7, if we confine our discussions to an appropriate class $D \subset A_3$, then we can show that appropriately modified maximum likelihood estimator is always third-order asymptotically efficient in the class $D$. In this section we shall develop our discussions by very general fashions. That is, we treat the ARMA$(p, q)$ process with multivariate unknown parameter $\boldsymbol{\theta}$.

## 2. Concepts of Higher Order Asymptotic Efficiency

In this section we shall explain a higher order asymptotic efficiency in the sense of highest probability concentration around the true value by the Edgeworth expansion. This concept of efficiency was introduced by Akahira and Takeuchi [3].

Now we consider the approach of Akahira and Takeuchi [3] whose argument proceeds as follows. Let $\mathbf{X}_T = (X_1,..., X_T)'$ denote a sequence of random variable forming a stochastic process, and possessing the probability measure $P_\theta^T(\cdot)$, where $\theta \in \Theta$, a subset of the real line. If an estimator $\hat{\theta}_T$ satisfies the equations

$$\lim_{T\to\infty} T^{(k-1)/2} |P_\theta^T\{\sqrt{T}(\hat{\theta}_T - \theta) \leqslant 0\} - \tfrac{1}{2}| = 0, \tag{2.1}$$

$$\lim_{T\to\infty} T^{(k-1)/2} |P_\theta^T\{\sqrt{T}(\hat{\theta}_T - \theta) \geqslant 0\} - \tfrac{1}{2}| = 0, \tag{2.2}$$

then $\hat{\theta}_T$ is called $k$th-order asymptotically median unbiased ($k$th-order AMU for short). We denote the set of $k$th-order AMU estimators by $A_k$. For $\hat{\theta}_T$ $k$th-order AMU,

$$F_0^+(x, \theta) + T^{-1/2}F_1^+(x, \theta) + \cdots + T^{-(k-1)/2}F_{k-1}^+(x, \theta)$$

and

$$F_0^-(x, \theta) + T^{-1/2}F_1^-(x, \theta) + \cdots + T^{-(k-1)/2}F_{k-1}^-(x, \theta)$$

are said to be the $k$th-order asymptotic distributions of $\sqrt{T}(\hat{\theta}_T - \theta)$ if

$$\begin{aligned}\lim_{T\to\infty} T^{(k-1)/2} | P_\theta^T\{\sqrt{T}(\hat{\theta}_T - \theta) \leqslant x\} - F_0^+(x, \theta) - T^{-1/2}F_1^+(x, \theta) \\ - \cdots - T^{-(k-1)/2}F_{k-1}^+(x, \theta)| = 0 \qquad \text{for all } x \geqslant 0, \qquad (2.3)\end{aligned}$$

$$\begin{aligned}\lim_{T\to\infty} T^{(k-1)/2} | P_\theta^T\{\sqrt{T}(\hat{\theta}_T - \theta) \leqslant x\} - F_0^-(x, \theta) - T^{-1/2}F_1^-(x, \theta) \\ - \cdots - T^{-(k-1)/2}F_{k-1}^-(x, \theta)| = 0 \qquad \text{for all } x < 0. \qquad (2.4)\end{aligned}$$

For $\theta_0 \in \Theta$, consider the problem of testing hypothesis $H^+ : \theta = \theta_0 + x/\sqrt{T}$ $(x > 0)$ against alternative $K : \theta = \theta_0$. We define

$$H_0^+(x, \theta_0) + T^{-1/2}H_1^+(x, \theta_0) + \cdots + T^{-(k-1)/2}H_{k-1}^+(x, \theta_0)$$

as follows

$$\begin{aligned}\sup_{\{A_T\}\in\Phi_X} \limsup_{T\to\infty} T^{(k-1)/2}\{P_{\theta_0}^T(A_T) - H_0^+(x, \theta_0) \\ - T^{-1/2}H_0^+(x, \theta_0) - \cdots - T^{-(k-1)/2}H_{k-1}^+(x, \theta_0)\} = 0, \quad (2.5)\end{aligned}$$

where $\Phi_X$ is the class of sets $A_T = \{\sqrt{T}(\tilde{\theta}_T - \theta) \leqslant x\}$ with $\tilde{\theta}_T$ $k$th order AMU. Then we have for $x > 0$,

$$\begin{aligned}P_{\theta_0 + x/\sqrt{T}}^T(A_T) &= P_{\theta_0 + x/\sqrt{T}}^T\{\sqrt{T}(\tilde{\theta}_T - \theta_0 - x/\sqrt{T}) \leqslant 0\} \\ &= \tfrac{1}{2} + o(T^{-(k-1)/2}).\end{aligned} \qquad (2.6)$$

By (2.3) and (2.5) we have

$$\begin{aligned}\limsup_{T\to\infty} T^{(k-1)/2}\{F_0^+(x, \theta_0) \\ + T^{-1/2}F_1^+(x, \theta_0) + \cdots + T^{-(k-1)/2}F_{k-1}^+(x, \theta_0) \\ - H_0^+(x, \theta_0) - T^{-1/2}H_1^+(x, \theta_0) - \cdots - T^{-(k-1)/2}H_{k-1}^+(x, \theta_0)\} \leqslant 0 \\ \text{for all } x > 0. \qquad (2.7)\end{aligned}$$

Also consider the problem of the testing hypothesis $H^- : \theta = \theta_0 + x/\sqrt{T}$ $(x < 0)$ against alternative $K : \theta = \theta_0$. Then we define

$$H_0^-(x, \theta_0) + T^{-1/2} H_1^-(x, \theta_0) + \cdots + T^{-(k-1)/2} H_{k-1}^-(x, \theta_0)$$

as follows

$$\inf_{\{A_T\} \in \Phi_X} \liminf_{T \to \infty} T^{(k-1)/2} \{ P_{\theta_0}^T(A_T) - H_0^-(x, \theta_0) - T^{-1/2} H_1^-(x, \theta_0) - \cdots - T^{-(k-1)/2} H_{k-1}^-(x, \theta_0) \} = 0. \tag{2.8}$$

In the same way as for the case $x > 0$, by (2.4) and (2.8) we have

$$\liminf_{T \to \infty} T^{(k-1)/2} \{ F_0^-(x, \theta_0) + T^{-1/2} F_1^-(x, \theta_0) + \cdots + T^{-(k-1)/2} F_{k-1}^-(x, \theta_0) - H_0^-(x, \theta_0) - T^{-1/2} H_1^-(x, \theta_0) - \cdots - T^{-(k-1)/2} H_{k-1}^-(x, \theta_0) \} \geqslant 0 \quad \text{for all } x < 0. \tag{2.9}$$

Thus we make the following definition.

DEFINITION 1 (Akahira and Takeuchi [3]). *A* $k$th-*order* AMU $\{\hat{\theta}_T\}$ *is called* $k$th-*order asymptotically efficient if for each* $\theta \in \Theta$,

$$\begin{aligned} \lim_{T \to \infty} & P_\theta^T \{ \sqrt{T}(\hat{\theta}_T - \theta) \leqslant x \} \\ &= H_0^+(x, \theta) + T^{-1/2} H_1^+(x, \theta) + \cdots + T^{-(k-1)/2} H_{k-1}^+(x, \theta) + o(T^{-(k-1)/2}) \quad \text{for all } x \geqslant 0, \\ &= H_0^-(x, \theta) + T^{-1/2} H_1^-(x, \theta) + \cdots + T^{-(k-1)/2} H_{k-1}^-(x, \theta) + o(T^{-(k-1)/2}) \quad \text{for all } x < 0. \end{aligned}$$

In the above discussion we can regard the bound distribution

$$H_0^+(x, \theta_0) + T^{-1/2} H_1^+(x, \theta_0) + \cdots + T^{-(k-1)/2} H_{k-1}^+(x, \theta_0)$$

as an approximation of the power function of the testing hypothesis $H^+ : \theta = \theta_0 + x/\sqrt{T}$ $(x > 0)$ against alternative $K : \theta = \theta_0$ at significance level $\frac{1}{2} + o(T^{-(k-1)/2})$. By the fundamental lemma of Neyman and Pearson this bound distribution can be given by deriving the asymptotic expansion of the likelihood ratio test which tests the null hypothesis $H^+ : \theta = \theta_0 + x/\sqrt{T}$ $(x > 0)$ against the alternative $K : \theta = \theta_0$ at significance level $\frac{1}{2} + o(T^{-(k-1)/2})$. In case of $x < 0$, we can proceed similarly.

## 3. Derivation of Asymptotic Expansion

As we saw in the previous section we use asymptotic expansions of the concerned statistics in our discussions of higher order efficiency. Thus, in this section we explain a derivation of the asymptotic expansion. Let $U_T=(u_1,..., u_p)'$ be a measurable function of a sample $X_1,..., X_T$. Suppose that all order of cumulants of $U_T$ exist and satisfy the following:

$$c_i=\mathrm{cum}(u_i)=T^{-1/2}c_i^{(1)}+T^{-1}c_i^{(2)}+o(T^{-1}), \tag{3.1}$$

$$c_{ij}=\mathrm{cum}(u_i, u_j)=c_{ij}^{(1)}+T^{-1/2}c_{ij}^{(2)}+T^{-1}c_{ij}^{(3)}+o(T^{-1}), \tag{3.2}$$

$$c_{ijk}=\mathrm{cum}(u_i, u_j, u_k)=T^{-1/2}c_{ijk}^{(1)}+T^{-1}c_{ijk}^{(2)}+o(T^{-1}), \tag{3.3}$$

$$c_{ijkm}=\mathrm{cum}(u_i, u_j, u_k, u_m)=T^{-1}c_{ijm}^{(1)}+o(T^{-1}), \tag{3.4}$$

$i, j, k, m=1,..., p$, and the $J$th-order cumulant satisfies

$$c_{i_1\cdots i_J}=\mathrm{cum}^{(J)}(u_{i_1},..., u_{i_J})=O(T^{-J/2+1}) \quad \text{for each } J\geqslant 5. \tag{3.5}$$

Then the characteristic function of $U_T$ is expressed as

$$\begin{aligned}
&\exp\Bigg\{\sum_i c_i(it_i)+\tfrac{1}{2}\sum_{i,j} c_{ij}(it_i)(it_j)+\tfrac{1}{6}\sum_{i,j,k} c_{ijk}(it_i)(it_j)(it_k)\\
&\quad+\tfrac{1}{24}\sum_{i,j,k,m} c_{ijkm}(it_i)(it_j)(it_k)(it_m)+\cdots\Bigg\}\\
&=\exp\left\{-\frac{1}{2}\sum_{i,j} c_{ij}^{(1)}t_it_j\right\}\Bigg[1+\sum_i\left(\frac{c_i^{(1)}}{\sqrt{T}}+\frac{c_i^{(2)}}{T}\right)(it_i)\\
&\quad+\frac{1}{2}\sum_{i,j}\left(\frac{c_{ij}^{(2)}}{\sqrt{T}}+\frac{c_{ij}^{(3)}}{T}+\frac{c_i^{(1)}c_j^{(1)}}{T}\right)(it_i)(it_j)\\
&\quad+\sum_{i,j,k}\left(\frac{c_{ijk}^{(1)}}{6\sqrt{T}}+\frac{c_{ijk}^{(2)}}{6T}+\frac{c_i^{(1)}c_{jk}^{(2)}}{2T}\right)(it_i)(it_j)(it_k)\\
&\quad+\sum_{i,j,k,m}\left(\frac{c_{ijkm}^{(1)}}{24T}+\frac{c_{ij}^{(2)}c_{km}^{(2)}}{8T}+\frac{c_i^{(1)}c_{jkm}^{(1)}}{6T}\right)(it_i)(it_j)(it_k)(it_m)\\
&\quad+\sum_{i,j,i',j',k'}\frac{c_{ij}^{(2)}c_{i'j'k'}^{(1)}}{12T}(it_i)(it_j)(it_{i'})(it_{j'})(it_{k'})\\
&\quad+\sum_{i,j,k,i',j',k'}\frac{c_{ijk}^{(1)}c_{i'j'k'}^{(1)}}{72T}(it_i)(it_j)(it_k)(it_{i'})(it_{j'})(it_{k'})\\
&\quad+o(T^{-1})\Bigg].
\end{aligned} \tag{3.6}$$

Inverting (3.6) by the Fourier inverse transform we have

$$
\begin{aligned}
&P(u_1 < y_1, \dots, u_p < y_p) \\
&= \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_p} N(y; \Omega) \Bigg[ 1 + \sum_i \left( \frac{c_i^{(1)}}{\sqrt{T}} + \frac{c_i^{(2)}}{T} \right) H_i(y) \\
&\quad + \frac{1}{2} \sum_{i,j} \left( \frac{c_{ij}^{(2)}}{\sqrt{T}} + \frac{c_{ij}^{(3)}}{T} + \frac{c_i^{(1)} c_j^{(1)}}{T} \right) H_{ij}(y) \\
&\quad + \sum_{i,j,k} \left( \frac{c_{ijk}^{(1)}}{6\sqrt{T}} + \frac{c_{ijk}^{(2)}}{6T} + \frac{c_i^{(1)} c_{jk}^{(2)}}{2T} \right) H_{ijk}(y) \\
&\quad + \sum_{i,j,k,m} \left( \frac{c_{ijkm}^{(1)}}{24T} + \frac{c_{ij}^{(2)} c_{km}^{(2)}}{8T} + \frac{c_i^{(1)} c_{jkm}^{(1)}}{6T} \right) H_{ijkm}(y) \\
&\quad + \sum_{i,j,i',j',k'} \frac{c_{ij}^{(2)} c_{i'j'k'}^{(1)}}{12T} H_{iji'j'k'}(y) \\
&\quad + \sum_{i,j,k,i',j',k'} \frac{c_{ijk}^{(1)} c_{i'j'k'}^{(1)}}{72T} H_{ijki'j'k'}(y) \Bigg] dy + o(T^{-1}),
\end{aligned}
\tag{3.7}
$$

where $y = (y_1, \dots, y_p)'$, $N(y; \Omega) = (2\pi)^{-p/2} |\Omega|^{-1/2} \exp(-\frac{1}{2} y' \Omega^{-1} y)$,

$$
H_{j_1 \cdots j_s}(y) = \frac{(-1)^s}{N(y; \Omega)} \frac{\partial^s}{\partial y_{j_1} \cdots \partial y_{j_s}} N(y; \Omega), \qquad \text{and} \qquad \Omega = (c_{ij}^{(1)}).
$$

In the special case of $p = 1$ and $c_{11}^{(1)} = 1$, we have

$$
\begin{aligned}
P(u_1 < y_1) = \Phi(y_1) - \phi(y_1) \Bigg[ & \frac{c_1^{(1)}}{\sqrt{T}} + \frac{c_1^{(2)}}{T} \\
& + \frac{1}{2} \left( \frac{c_{11}^{(2)}}{\sqrt{T}} + \frac{c_{11}^{(3)}}{T} + \frac{c_1^{(1)} c_1^{(1)}}{T} \right) y_1 \\
& + \left( \frac{c_{111}^{(1)}}{6\sqrt{T}} + \frac{c_{111}^{(2)}}{6T} + \frac{c_1^{(1)} c_{11}^{(2)}}{2T} \right) (y_1^2 - 1) \\
& + \left( \frac{c_{1111}^{(1)}}{24T} + \frac{c_{11}^{(2)} c_{11}^{(2)}}{8T} + \frac{c_1^{(1)} c_{111}^{(1)}}{6T} \right) (y_1^3 - 3y_1) \\
& + \frac{c_{11}^{(2)} c_{111}^{(1)}}{12T} (y_1^4 - 6y_1^2 + 3) \\
& + \frac{c_{111}^{(1)} c_{111}^{(1)}}{72T} (y_1^5 - 10y_1^3 + 15y_1) \Bigg] + o(T^{-1}),
\end{aligned}
\tag{3.8}
$$

where $\Phi(y) = \int_{-\infty}^{y} \phi(t)\, dt$, $\phi(t) = (1/\sqrt{2\pi}) \exp -(t^2/2)$. Fang and Krishnaiah (1982) gave the Edgeworth expansion for a function of sum of independent non-normal random vectors.

## 4. Evaluation Method of Asymptotic Moments

To derive the asymptotic expansions for some basic statistics, as we saw in Section 3, it is required to evaluate the asymptotic cumulants or moments of the statistics.

Now we introduce $D_\Delta$ and $D_{\mathrm{ARMA}}$, the spaces of functions on $[-\pi, \pi]$;

$$D_\Delta = \left\{ f : f(\lambda) = \sum_{u=-\infty}^{\infty} a(u) \exp(-iu\lambda),\ a(u) = a(-u), \right.$$

$$\left. \sum_{u=-\infty}^{\infty} |u|^2 |a(u)| < \infty \right\},$$

$$D_{\mathrm{ARMA}} = \{ f : f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{s} \alpha_j e^{ij\lambda} \right|^2 \Big/ \left| \sum_{j=0}^{r} \beta_j e^{ij\lambda} \right|^2, \qquad (\sigma^2 > 0),$$

for some positive integers $r$ and $s$, where $A(z) = \sum_{j=0}^{s} \alpha_j z^j$ and $B(z) = \sum_{j=0}^{r} \beta_j z^j$ are bounded away from zero for $|z| \leqslant 1\}$.

We set down the following assumptions.

Assumption 1. *The process $\{X_t;\ t = 0, \pm 1, \pm 2,...,\}$ is a Gaussian stationary process with the spectral density $f_{\boldsymbol\theta}(\lambda) \in D_{\mathrm{ARMA}}$, $\boldsymbol\theta \in \Theta \subset R^p$, and mean* 0.

Assumption 2. *The spectral density $f_{\boldsymbol\theta}(\lambda)$ is continuously five times differentiable with respect to $\boldsymbol\theta$, and the derivatives $\partial f_{\boldsymbol\theta}/\partial\theta_j$, $\partial^2 f_{\boldsymbol\theta}/\partial\theta_j\partial\theta_k$,..., $\partial^5 f_{\boldsymbol\theta}/\partial\theta_j\partial\theta_k\partial\theta_m\partial\theta_l\partial\theta_n$ $(j, k, m, l, n = 1,..., p)$* belong to $D_\Delta$.

Assumption 3. *If $\boldsymbol\theta \neq \boldsymbol\theta^*$, then $f_{\boldsymbol\theta} \neq f_{\boldsymbol\theta^*}$ on a set of positive Lebesgue measure.*

Assumption 4. *The matrix*

$$I(\boldsymbol\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \frac{\partial}{\partial \boldsymbol\theta} \log f_{\boldsymbol\theta}(\lambda) \right\} \left\{ \frac{\partial}{\partial \boldsymbol\theta'} \log f_{\boldsymbol\theta}(\lambda) \right\} d\lambda$$

*is positive definite for all $\boldsymbol\theta \in \Theta$.*

Suppose that a stretch, $\mathbf{X}_T = (X_1,..., X_T)'$ of the series $\{X_t\}$ is available. Let $\Sigma_T$ be the covariance matrix of $\mathbf{X}_T$. The likelihood function based on $\mathbf{X}_T$ is given by

$$l_T(\boldsymbol\theta) = (2\pi)^{-T/2} |\Sigma_T|^{-1/2} \exp\{ -\tfrac{1}{2} \mathbf{X}'_T \Sigma_T^{-1} \mathbf{X}_T \}.$$

Let

$$Z_i = \frac{1}{\sqrt{T}} \frac{\partial \log l_T(\boldsymbol{\theta})}{\partial \theta_i}, \tag{4.1}$$

$$Z_{ij} = \frac{1}{\sqrt{T}} \left\{ \frac{\partial^2 \log l_T(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} - E_{\boldsymbol{\theta}} \frac{\partial^2 \log l_T(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right\}, \tag{4.2}$$

and

$$Z_{ijk} = \frac{1}{\sqrt{T}} \left\{ \frac{\partial^3 \log l_T(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} - E_{\boldsymbol{\theta}} \frac{\partial^3 \log l_T(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} \right\}. \tag{4.3}$$

Hereafter we shall treat statistics which are approximated by simple functions of $Z_i$, $Z_{ij}$, and $Z_{ijk}$. To give their asymptotic expansions, we evaluate the asymptotic cumulants (moments) of $Z_i$, $Z_{ij}$, and $Z_{ijk}$. Here we can see that

$$\frac{\partial \log l_T(\boldsymbol{\theta})}{\partial \theta_i} = \frac{1}{2} \boldsymbol{X}'_T \Sigma_T^{-1} \Sigma_T^{(i)} \Sigma_T^{-1} \mathbf{X}_T - \frac{1}{2} \operatorname{tr} \Sigma_T^{-1} \Sigma_T^{(i)}, \tag{4.4}$$

$$\begin{aligned} \frac{\partial^2 \log l_T(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} = & -\frac{1}{2} \mathbf{X}'_T \Sigma_T^{-1} \{ \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(i)} \\ & + \Sigma_T^{(i)} \Sigma_T^{-1} \Sigma_T^{(j)} - \Sigma_T^{(i,j)} \} \Sigma_T^{-1} \mathbf{X}_T \\ & - \tfrac{1}{2} \operatorname{tr} \{ \Sigma_T^{-1} \Sigma_T^{(i,j)} - \Sigma_T^{-1} \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(i)} \}, \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} & \frac{\partial^3 \log l_T(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} \\ & = \tfrac{1}{2} \boldsymbol{X}'_T \Sigma_T^{-1} \{ \Sigma_T^{(k)} \Sigma_T^{-1} \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(i)} + \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(k)} \Sigma_T^{-1} \Sigma_T^{(i)} \\ & \quad + \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(i)} \Sigma_T^{-1} \Sigma_T^{(k)} + \Sigma_T^{(k)} \Sigma_T^{-1} \Sigma_T^{(i)} \Sigma_T^{-1} \Sigma_T^{(j)} \\ & \quad + \Sigma_T^{(i)} \Sigma_T^{-1} \Sigma_T^{(k)} \Sigma_T^{-1} \Sigma_T^{(j)} \\ & \quad + \Sigma_T^{(i)} \Sigma_T^{-1} \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(k)} - \Sigma_T^{(j,k)} \Sigma_T^{-1} \Sigma_T^{(i)} \\ & \quad - \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(i,k)} - \Sigma_T^{(i,k)} \Sigma_T^{-1} \Sigma_T^{(j)} \\ & \quad - \Sigma_T^{(i)} \Sigma_T^{-1} \Sigma_T^{(j,k)} - \Sigma_T^{(k)} \Sigma_T^{-1} \Sigma_T^{(i,j)} \\ & \quad - \Sigma_T^{(i,j)} \Sigma_T^{-1} \Sigma_T^{(k)} + \Sigma_T^{(i,j,k)} \} \Sigma_T^{-1} \mathbf{X}_T \\ & \quad - \tfrac{1}{2} \operatorname{tr} \{ \Sigma_T^{-1} \Sigma_T^{(k)} \Sigma_T^{-1} \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(i)} \\ & \quad + \Sigma_T^{-1} \Sigma_T^{(j)} \Sigma_T^{-1} \Sigma_T^{(k)} \Sigma_T^{-1} \Sigma_T^{(i)} - \Sigma_T^{-1} \Sigma_T^{(k)} \Sigma_T^{-1} \Sigma_T^{(i,j)} \\ & \quad - \Sigma_T^{-1} \Sigma_T^{(j,k)} \Sigma_T^{-1} \Sigma_T^{(i)} - \Sigma_T^{-1} \Sigma_T^{(j)} \Sigma_T^{-1)} \Sigma_T^{(i,k)} + \Sigma_T^{-1} \Sigma_T^{(i,j,k)} \}, \end{aligned} \tag{4.6}$$

where $\Sigma_T^{(i)}$, $\Sigma_T^{(i,j)}$, and $\Sigma_T^{(i,j,k)}$ are the $T \times T$—Toeplitz type matrices whose $(l, m)$th elements are given by

$$\int_{-\pi}^{\pi} e^{i(l-m)\lambda} \frac{\partial}{\partial \theta_i} f_{\boldsymbol{\theta}}(\lambda)\, d\lambda, \qquad \int_{-\pi}^{\pi} e^{i(l-m)\lambda} \frac{\partial^2}{\partial \theta_i \partial \theta_j} f_{\boldsymbol{\theta}}(\lambda)\, d\lambda$$

and

$$\int_{-\pi}^{\pi} e^{i(l-m)\lambda} \frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} f_{\boldsymbol{\theta}}(\lambda)\, d\lambda, \text{ respectively.}$$

To evaluate the asymptotic moments of $Z_i$, $Z_{ij}$, and $Z_{ijk}$, we need the following lemmas. The first one is due to Magnus and Neudecker [14].

LEMMA 1. *Let $A$, $B$, and $C$ be symmetric matrices of order $T$ (non-random); then*

(i) $E(\mathbf{X}_T' A \mathbf{X}_T \cdot \mathbf{X}_T' B \mathbf{X}_X) = (\operatorname{tr} A\Sigma_T)(\operatorname{tr} B\Sigma_T) + 2 \operatorname{tr} A\Sigma_T B\Sigma_T$,

(ii) $E(\mathbf{X}_T' A \mathbf{X}_T \cdot \mathbf{X}_T' B \mathbf{X}_T \cdot \mathbf{X}_T' C \mathbf{X}_T) = (\operatorname{tr} A\Sigma_T)(\operatorname{tr} B\Sigma_T)(\operatorname{tr} C\Sigma_T)$

$+ 2[(\operatorname{tr} A\Sigma_T)(\operatorname{tr} B\Sigma_T C\Sigma_T) + (\operatorname{tr} B\Sigma_T)(\operatorname{tr} A\Sigma_T C\Sigma_T)$

$+ (\operatorname{tr} C\Sigma_T)(\operatorname{tr} A\Sigma_T B\Sigma_T)] + 8 \operatorname{tr} A\Sigma_T B\Sigma_T C\Sigma_T$.

The second one is due to Taniguchi [15].

LEMMA 2. *Suppose that $f_1(\lambda),..., f_s(\lambda) \in D_\Lambda$, $g_1(\lambda),..., g_s(\lambda) \in D_{\mathrm{ARMA}}$. We define $\Gamma_1,..., \Gamma_s$, $\Lambda_1,..., \Lambda_s$, the $T \times T$-Toeplitz type matrices, by*

$$\Gamma_j = \left( \int_{-\pi}^{\pi} e^{i(k-l)\lambda} f_j(\lambda)\, d\lambda \right), \qquad \Lambda_j = \left( \int_{-\pi}^{\pi} e^{i(k-l)\lambda} g_j(\lambda)\, d\lambda \right)$$

*$(k, l = 1,..., T, j = 1,..., s)$. Then*

$$T^{-1} \operatorname{tr} \Gamma_1 \Lambda_1^{-1} \Gamma_2 \Lambda_2^{-1} \cdots \Gamma_s \Lambda_s^{-1}$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_1(\lambda) \cdots f_s(\lambda)\, g_1(\lambda)^{-1} \cdots g_s(\lambda)^{-1}\, d\lambda + O(T^{-1}).$$

Using Lemmas 1 and 2, it is not difficult to show the following lemma.

LEMMA 3. *Under Assumptions 1–4, we have*

$$E(Z_i Z_j) = I_{ij} + O(T^{-1}), \tag{4.7}$$

$$E(Z_i Z_{jk}) = J_{ijk} + O(T^{-1}), \tag{4.8}$$

$$E(Z_i Z_j Z_k) = \frac{1}{\sqrt{T}} K_{ijk} + O(T^{-3/2}), \tag{4.9}$$

$$E(Z_i Z_{jkm}) = L_{ijkm} + O(T^{-1}), \tag{4.10}$$

$$\mathrm{Cov}(Z_{ij}, Z_{km}) = M_{ijkm} + O(T^{-1}), \tag{4.11}$$

$$E(Z_i Z_j Z_{km}) = \frac{1}{\sqrt{T}} N_{ijkm} + O(T^{-3/2}), \tag{4.12}$$

$$\mathrm{cum}\{Z_i, Z_j, Z_k, Z_m\} = \frac{1}{T} H_{ijkm} + O(T^{-2}), \tag{4.13}$$

*where*

$$I_{ij} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_j} \{f_{\boldsymbol{\theta}}(\lambda)\} f_{\boldsymbol{\theta}}(\lambda)^{-2}\, d\lambda, \tag{4.14}$$

$$\begin{aligned} J_{ijk} = & -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_j} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_k} \{f_{\boldsymbol{\theta}}(\lambda)\} f_{\boldsymbol{\theta}}(\lambda)^{-3}\, d\lambda \\ & + \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial^2}{\partial\theta_j \partial\theta_k} \{f_{\boldsymbol{\theta}}(\lambda)\} f_{\boldsymbol{\theta}}(\lambda)^{-2}\, d\lambda, \end{aligned} \tag{4.15}$$

$$K_{ijk} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_j} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_k} \{f_{\boldsymbol{\theta}}(\lambda)\} f_{\boldsymbol{\theta}}(\lambda)^{-3}\, d\lambda, \tag{4.16}$$

$$\begin{aligned} L_{ijkm} = & \frac{3}{2\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_j} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_k} \{f_{\boldsymbol{\theta}}(\lambda)\} \\ & \times \frac{\partial}{\partial\theta_m} \{f_{\boldsymbol{\theta}}(\lambda)\} f_{\boldsymbol{\theta}}(\lambda)^{-4}\, d\lambda \\ & - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \{f_{\boldsymbol{\theta}}(\lambda)\} \Bigg[ \frac{\partial^2}{\partial\theta_k \partial\theta_m} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_j} \{f_{\boldsymbol{\theta}}(\lambda)\} \\ & + \frac{\partial^2}{\partial\theta_j \partial\theta_k} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_m} \{f_{\boldsymbol{\theta}}(\lambda)\} \\ & + \frac{\partial^2}{\partial\theta_j \partial\theta_m} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_k} \{f_{\boldsymbol{\theta}}(\lambda)\} \Bigg] f_{\boldsymbol{\theta}}(\lambda)^{-3}\, d\lambda \\ & + \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial^3}{\partial\theta_j \partial\theta_k \partial\theta_m} \{f_{\boldsymbol{\theta}}(\lambda)\} f_{\boldsymbol{\theta}}(\lambda)^{-2}\, d\lambda, \end{aligned} \tag{4.17}$$

$$\begin{aligned} M_{ijkm} = & \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_j} \{f_{\boldsymbol{\theta}}(\lambda)\} \frac{\partial}{\partial\theta_k} \{f_{\boldsymbol{\theta}}(\lambda)\} \\ & \times \frac{\partial}{\partial\theta_m} \{f_{\boldsymbol{\theta}}(\lambda)\} f_{\boldsymbol{\theta}}(\lambda)^{-4}\, d\lambda \end{aligned}$$

$$-\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{\partial}{\partial\theta_i}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial}{\partial\theta_j}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial^2}{\partial\theta_k\partial\theta_m}\{f_{\boldsymbol\theta}(\lambda)\}\right.$$

$$\left.+\frac{\partial}{\partial\theta_k}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial}{\partial\theta_m}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\{f_{\boldsymbol\theta}(\lambda)\}\right]f_{\boldsymbol\theta}(\lambda)^{-3}\,d\lambda$$

$$+\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial^2}{\partial\theta_k\partial\theta_m}\{f_{\boldsymbol\theta}(\lambda)\}f_{\boldsymbol\theta}(\lambda)^{-2}\,d\lambda, \tag{4.18}$$

$$N_{ijkm}=-\frac{1}{\pi}\int_{-\pi}^{\pi}\frac{\partial}{\partial\theta_i}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial}{\partial\theta_j}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial}{\partial\theta_k}\{f_{\boldsymbol\theta}(\lambda)\}$$

$$\times\frac{\partial}{\partial\theta_m}\{f_{\boldsymbol\theta}(\lambda)\}f_{\boldsymbol\theta}(\lambda)^{-4}\,d\lambda$$

$$+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial}{\partial\theta_i}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial}{\partial\theta_j}\{f_{\boldsymbol\theta}(\lambda)\}$$

$$\times\frac{\partial^2}{\partial\theta_k\,\partial\theta_m}\{f_{\boldsymbol\theta}(\lambda)\}f_{\boldsymbol\theta}(\lambda)^{-3}\,d\lambda, \tag{4.19}$$

$$H_{ijkm}=\frac{3}{2\pi}\int_{-\pi}^{\pi}\frac{\partial}{\partial\theta_i}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial}{\partial\theta_j}\{f_{\boldsymbol\theta}(\lambda)\}$$

$$\times\frac{\partial}{\partial\theta_k}\{f_{\boldsymbol\theta}(\lambda)\}\frac{\partial}{\partial\theta_m}\{f_{\theta}(\lambda)\}f_{\boldsymbol\theta}(\lambda)^{-4}\,d\lambda. \tag{4.20}$$

Henceforth, if $\boldsymbol\theta$ is scalar, we use $I$, $J$, $K$, etc. (or $I(\theta)$, $J(\theta)$, $K(\theta)$, etc.) instead of $I_{ij}$, $J_{ijk}$, $K_{ijk}$, etc. for simplicity. In the subsequent sections we need the following lemmas to prove the validity of Edgeworth expansions. Put $B_j=\Gamma_1^{(j)}\Lambda_1^{(j)^{-1}}\cdots\Gamma_{s(j)}^{(j)}\Lambda_{s(j)}^{(j)^{-1}}$, where $\{\Gamma_l^{(j)}\}$ and $\{\Lambda_l^{(j)}\}$ are the same type of matrices as $\Gamma_l$ and $\Lambda_l$, respectively, given in Lemma 2. Then we have

LEMMA 4. *Under Assumptions* 1–4, *for every integer* $\eta\geqslant 1$, *the orders of the following* $2\eta$th *cumulant and moment are given by*

$$C_{2\eta}^T=\operatorname{cum}\left\{\frac{1}{\sqrt{T}}(\mathbf{X}'_T B_1\mathbf{X}_T-E\mathbf{X}'_T B_1\mathbf{X}_T),\ldots,\frac{1}{\sqrt{T}}(\mathbf{X}'_T B_{2\eta}\mathbf{X}_T-E\mathbf{X}'_T B_{2\eta}\mathbf{X}_T)\right\}$$

$$=O(T^{-\eta+1}), \tag{4.21}$$

$$M_{2\eta}^T=E\left[\frac{1}{\sqrt{T}}(\mathbf{X}'_T B_1\mathbf{X}_T-E\mathbf{X}'_T B_1\mathbf{X}_T)\right]^{2\eta}$$

$$=O(1), \tag{4.22}$$

*respectively.*

*Proof.* We can represent

$$C_{2\eta}^{T}=T^{-\eta}\,\mathrm{cum}\{\mathbf{z}_T'\Sigma_T^{-1/2}B_1\Sigma_T^{-1/2}\mathbf{z}_T,..., \mathbf{z}_T'\Sigma_T^{-1/2}B_{2\eta}\Sigma_T^{-1/2}\mathbf{z}_T\}, \quad (4.23)$$

where $\mathbf{z}_T'=(z_1,..., z_T)$ is a random vector distributed as $N(\mathbf{O}_T, I_T)$, ($I_T$ is the $T\times T$-identity matrix). Denote the $(i,j)$th component of $\Sigma_T^{-1/2}B_l\Sigma_T^{-1/2}$ by $b_{ij}^{(l)}$. Then, using the fundamental properties of cumulant, we have

$$\begin{aligned} C_{2\eta}^{T}=T^{-\eta}&\sum_{i_1=1}^{T}\sum_{j_1=1}^{T}\cdots\sum_{i_{2\eta}=1}^{T}\sum_{j_{2\eta}=1}^{T} b_{i_1j_1}^{(1)}\cdots b_{i_{2\eta}j_{2\eta}}^{(2\eta)}\\ &\times\mathrm{cum}\{z_{i_1}z_{j_1},..., z_{i_{2\eta}}z_{j_{2\eta}}\}, \end{aligned} \quad (4.24)$$

(see Theorem 2.3.1 of Brillinger [6]). Noting Theorem 2.3.2 of Brillinger [6], it is not difficult to show

$$C_{2\eta}^{T}=O\left\{T^{-\eta}\,\mathrm{tr}\prod_{l=1}^{2\eta}(\Sigma_T^{-1}B_l)\right\}.$$

By Lemma 2, we have $\mathrm{tr}\prod_{l=1}^{2\eta}(\Sigma_T^{-1}B_l)=O(T)$, which completes the proof of (4.21). As for (4.22), making use of the identity (see Brillinger [5]),

$$E(Y_1Y_2\cdots Y_{2\eta})=\sum_{\nu}\mathrm{cum}\{Y_j\,(j\in\nu_1)\}\cdots\mathrm{cum}\{Y_j\,(j\in\nu_p)\},$$

where the summation is over all partitions $(\nu_1, \nu_2,..., \nu_p)$ $(p=1,..., 2\eta)$ of integers $1,..., 2\eta$, and noting (4.21), we have (4.22). ∎

LEMMA 5. *Under Assumptions* 1–4, *for every integer* $\eta\geq 1$, *we have*

$$P\left\{\frac{1}{T}|\mathbf{X}_T'B_1\mathbf{X}_T|>\rho_TT^{1/2}\right\}=o(T^{-\eta}), \quad (4.25)$$

*for some sequence* $\rho_T\to 0$, $\rho_TT^{1/2}\to\infty$ *as* $T\to\infty$.

*Proof.* By Tchebychev's inequality, we have

$$P\left\{\frac{1}{\sqrt{T}}|\mathbf{X}_T'B_1\mathbf{X}_T-E\mathbf{X}_T'B_1\mathbf{X}_T|>a\right\}\leqslant M_{2\eta}^{T}/a^{2\eta}, \quad (4.26)$$

for any $a>0$. Then we have

$$P\left\{\frac{1}{T}|X_T'B_1\mathbf{X}_T|>\frac{a}{\sqrt{T}}+\frac{|E\mathbf{X}_T'B_1\mathbf{X}_T|}{T}\right\}\leqslant M_{2\eta}^{T}/a^{2\eta}. \quad (4.27)$$

Noting that $(1/T)\,E\mathbf{X}_T'B_1\mathbf{X}_T=O(1)$ and $M_{2\eta}^{T}<\infty$, and choosing $a=T\rho_T$ in (4.27), we have the desired result. ∎

## 5. Third-Order Bound Distribution for the Class of Third-Order AMU Estimators

In this section we shall give the third-order bound distribution given in (2.5) for third-order AMU estimators. Let $\mathbf{X}_T = (X_1, ..., X_T)'$ be an observed stretch of the series $\{X_t\}$. We denote the covariance matrix of $\mathbf{X}_T$ by $\Sigma_T$. The log-likelihood function based on $\mathbf{X}_T$ is given by

$$G(\theta) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\log|\Sigma_T| - \frac{1}{2}\mathbf{X}_T'\Sigma_T^{-1}\mathbf{X}_T. \tag{5.1}$$

Consider the problem of testing hypothesis $H: \theta = \theta_0 + x/\sqrt{T}$ $(x > 0)$ against alternative $K: \theta = \theta_0$. Let $LR = G(\theta_0) - G(\theta_1)$, where $\theta_1 = \theta_0 + x/\sqrt{T}$. To give the bound distribution (see (2.5)), we shall derive the Edgeworth expansion of $LR$. Since the spectral density $f_\theta(\lambda)$ is continuously five times differentiable we get

$$\begin{aligned} LR = &-\frac{x}{\sqrt{T}}\left\{\frac{\partial}{\partial\theta}G(\theta)\right\}_{\theta_0} - \frac{x^2}{2T}\left\{\frac{\partial^2}{\partial\theta^2}G(\theta)\right\}_{\theta_0} - \frac{x^3}{6T\sqrt{T}}\left\{\frac{\partial^3}{\partial\theta^3}G(\theta)\right\}_{\theta_0} \\ &-\frac{x^4}{24T^2}\left\{\frac{\partial^4}{\partial\theta^4}G(\theta)\right\}_{\theta_0} - \frac{x^5}{120T^2\sqrt{T}}\left\{\frac{\partial^5}{\partial\theta^5}G(\theta)\right\}_{\theta'}, \end{aligned}$$

where $\theta_0 \leqslant \theta' \leqslant \theta_1$. Here we can express as $(\partial^5/\partial\theta^5)\, G(\theta) = \mathbf{X}_T' B_1 \mathbf{X}_T + \operatorname{tr} B_2$, where $B_1$ and $B_2$ are the same type matrices given in Lemmas 4 and 5. Using Lemmas 2 and 5, it is easy to show that

$$P\left\{\frac{1}{T}\left|\frac{\partial^5}{\partial\theta^5}G(\theta)\right| > \rho_T T^{1/2}\right\} = o(T^{-1}). \tag{5.2}$$

The following lemma is essentially due to Chibisov [7].

**Lemma 6.** *Let $Y^{(T)}$ be a random variable which has the stochastic expansion*

$$Y^{(T)} = Y_r^{(T)} + T^{-(r+1)/2}\xi_T,$$

*where $Y_r^{(T)}$ has the valid Edgeworth expansion up to order $T^{-r/2}$ and $\xi_T$ satisfies*

$$P\{|\xi_T| > \rho_T\sqrt{T}\} = o(T^{-r/2}),$$

*where $\rho_T \to 0$, $\rho_T T^{1/2} \to \infty$ as $T \to \infty$. Then $Y^{(T)}$ has the same Edgeworth expansion as that of $Y_r^{(T)}$ up to order $T^{-r/2}$.*

Noting (5.2) and Lemma 6, in order to derive the Edgeworth expansion for $LR$ up to order $T^{-1}$, we have only to derive that of

$$\widetilde{LR} = -\frac{x}{\sqrt{T}}\left\{\frac{\partial}{\partial\theta}G(\theta)\right\}_{\theta_0} - \frac{x^2}{2T}\left\{\frac{\partial^2}{\partial\theta_2}G(\theta)\right\}_{\theta_0} - \frac{x^3}{6T\sqrt{T}}\left\{\frac{\partial^3}{\partial\theta^3}G(\theta)\right\}_{\theta_0}$$
$$-\frac{x^4}{24T^2}\left\{\frac{\partial^4}{\partial\theta^4}G(\theta)\right\}_{\theta_0}.$$

Thus we evaluate the cumulants (moments) of $\widetilde{LR}$ under $\theta=\theta_0$ and $\theta=\theta_1$. We write

$$E_\theta\left\{\frac{1}{\sqrt{T}}\frac{\partial}{\partial\theta}G(\theta)\right\}^2 = I(\theta)+\Delta(\theta)/T+o(T^{-1}), \tag{5.3}$$

where $I(\theta)=(1/4\pi)\int_{-\pi}^{\pi}\{(\partial/\partial\theta)f_\theta(\lambda)\}^2\ f_\theta(\lambda)^{-2}\,d\lambda$, and $\Delta(\theta)$ will be explicitly evaluated in the case of ARMA(1, 1). Using Lemmas 1, 2, and 3, it is not difficult to show

$$E_{\theta_0}(\widetilde{LR}) = \frac{x^2}{2}I(\theta_0)+\frac{x^3}{6\sqrt{T}}(3J(\theta_0)+K(\theta_0))+\frac{x^2}{2T}\Delta(\theta_0)$$
$$+\frac{x^4}{24T}(4L(\theta_0)+3M(\theta_0)+6N(\theta_0)+H(\theta_0))+o(T^{-1}), \tag{5.4}$$

$$E_{\theta_0}(\widetilde{LR}-E_{\theta_0}\widetilde{LR})^2 = x^2I(\theta_0)+\frac{x^3}{\sqrt{T}}J(\theta_0)+\frac{x^4}{4T}M(\theta_0)+\frac{x^4}{3T}L(\theta_0)$$
$$+\frac{x^2}{T}\Delta(\theta_0)+o(T^{-1}), \tag{5.5}$$

$$\mathrm{cum}_{\theta_0}\{\widetilde{LR},\widetilde{LR},\widetilde{LR}\} = -\frac{x^3}{\sqrt{T}}K(\theta_0)-\frac{3x^4}{2T}N(\theta_0)+o(T^{-1}), \tag{5.6}$$

$$\mathrm{cum}_{\theta_0}\{\widetilde{LR},\widetilde{LR},\widetilde{LR},\widetilde{LR}\} = \frac{x^4}{T}H(\theta_0)+o(T^{-1}). \tag{5.7}$$

Remembering Lemma 4, we can see that the $J$th order cumulant satisfies

$$\mathrm{cum}_{\theta_0}^{(J)}\{\widetilde{LR},\ldots,\widetilde{LR}\} = O(T^{-J/2+1}) \qquad \text{for } J\geq 5. \tag{5.8}$$

Putting $V_T=\{\widetilde{LR}-E_{\theta_1}(\widetilde{LR})\}/\{x\sqrt{I_T}\}$, where $I_T=I+\Delta/T$, similarly we have

$$\mathrm{Var}_{\theta_1}(V_T) = 1 + \frac{1}{\sqrt{T}} b_1 + \frac{1}{T} b_2 + o(T^{-1}), \tag{5.9}$$

$$\mathrm{cum}_{\theta_1}\{V_T, V_T, V_T\} = \frac{1}{\sqrt{T}} c_1 + \frac{1}{T} c_2 + o(T^{-1}), \tag{5.10}$$

$$\mathrm{cum}_{\theta_1}\{V_T, V_T, V_T, V_T\} = \frac{1}{T} d_1 + o(T^{-1}), \tag{5.11}$$

$$\mathrm{cum}_{\theta_1}^{(J)}\{V_T, \ldots, V_T\} = O(T^{-J/2+1}), \qquad J \geqslant 5, \tag{5.12}$$

where $b_1 = (x/I)(J+K)$, $b_2 = (x^2/12I)(4L + 3M + 18N + 6H)$, $c_1 = -K/I^{3/2}$, $c_2 = -(x/2I^{3/2})(3N + 2H)$, $d_1 = H/I^2$. Remembering (3.8), we get the following Edgeworth expansion:

$$\begin{aligned} P_{\theta_1}^T(V_T \leqslant a) = \Phi(a) - \phi(a)\Bigg[&\frac{1}{2}\left(\frac{b_1}{\sqrt{T}} + \frac{b_2}{T}\right) a \\ &+ \left(\frac{c_1}{6\sqrt{T}} + \frac{c_2}{6T}\right)(a^2 - 1) \\ &+ \left(\frac{d_1}{24T} + \frac{b_1^2}{8T}\right)(a^3 - 3a) + \frac{b_1 c_1}{12T}(a^4 - 6a^2 + 3) \\ &+ \frac{c_1^2}{72T}(a^5 - 10a^3 + 15a)\Bigg] + o(T^{-1}). \end{aligned} \tag{5.13}$$

Noting $\Phi(a) = \frac{1}{2} + a\phi(a) - (a^3/2)\,\phi(a) + \ldots$, and if we put $a = -c_1/6\sqrt{T} - c_2/6T + b_1 c_1/6T$ in (5.18), it is easy to show

$$P_{\theta_1}^T(V_T \leqslant a) = \tfrac{1}{2} + o(T^{-1}), \tag{5.14}$$

$$P_{\theta_1}^T(V_T \geqslant a) = \tfrac{1}{2} + o(T^{-1}). \tag{5.15}$$

Putting $W_T = -\{V_T - a - x\sqrt{I_T}\}$ and $x' = x\sqrt{I_T}$, similarly we get the following Edgeworth expansion

$$\begin{aligned} P_{\theta_0}^T(V_T \geqslant a) &= P_{\theta_0}^T(W_T \leqslant x') \\ &= \Phi(x') - \phi(x')\Bigg\{\frac{\beta_3}{6\sqrt{T}} + \frac{\beta_3}{6\sqrt{T}}(x'^2 - 1) + \left(\frac{\beta_2}{2T} + \frac{\beta_3^2}{72T}\right) x' \\ &\quad + \left(\frac{\beta_4}{24T} + \frac{\beta_3^2}{36T}\right)(x'^3 - 3x') \\ &\quad + \frac{\beta_3^2}{72T}(x'^5 - 10x'^3 + 15x')\Bigg\} + o(T^{-1}), \end{aligned} \tag{5.16}$$

where $\beta_3 = -(3J+2K)/I^{3/2}$, $\beta_4 = 3\beta_3^2 - (4L+3M+12N+3H)/I^2$,

$$\beta_2 = \frac{17}{36}\beta_3^2 - \frac{K^2}{18I^3} - \frac{12L+9M+36N+8H}{12I^2}.$$

If $\{\hat{\theta}_T\}$ is third order AMU, then remembering (2.7), the fundamental lemma of Neyman and Pearson, and $\Phi(x') = \Phi(x\sqrt{I}) + (x\Delta/2T\sqrt{I})\phi(x\sqrt{I}) + o(T^{-1})$, we have

THEOREM 1. *For any $\hat{\theta}_T \in A_3$, we have*

$$\limsup_{T\to\infty} T[P_{\theta_0}^T\{\sqrt{TI}(\hat{\theta}_T - \theta_0) \leqslant y\} - F_{\theta_0}^{(3)}(y)] \leqslant 0 \qquad \textit{for } y > 0, \tag{5.17}$$

*where*

$$F_{\theta_0}^{(3)}(y) = \Phi(y) - \phi(y)\left\{\frac{\beta_3}{6\sqrt{T}} + \frac{\beta_3}{6\sqrt{T}}(y^2-1) + \left(\frac{\beta_2}{2T} + \frac{\beta_3}{72T} - \frac{\Delta}{2IT}\right)y \right.$$
$$\left. + \left(\frac{\beta_4}{24T} + \frac{\beta_3^2}{36T}\right)(y^3 - 3y) + \frac{\beta_3^2}{72T}(y^5 - 10y^3 + 15y)\right\}. \tag{5.18}$$

*For $y < 0$, similarly we have*

$$\liminf_{T\to\infty} T[P_{\theta_0}^T\{\sqrt{TI}(\hat{\theta}_T - \theta_0) \leqslant y\} - F_{\theta_0}^{(3)}(y)] \geqslant 0. \tag{5.19}$$

Now we seek the bound distribution $F_\theta^{(3)}(y)$ for concrete parameterization of the spectral density. Calculations for $I$, $J$, $K$, etc, can be done by the residue theorem (see [15]). However direct calculations for $\Delta(\theta)$ are very troublesome (i.e., $T^{-1}$-order term of $T^{-1}\mathrm{tr}\, \Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\dot{\Sigma}_T$). So we make the following device. The first part of the following lemma is given by differencing log det $\Sigma_T$ twice. The second part is essentially due to Galbraith and Galbraith [11].

LEMMA 7. *Suppose that the spectral density $f_\theta(\lambda)$ of $\{X_t\}$ is given by*

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi}\frac{|1-\beta e^{i\lambda}|^2}{|1-\alpha e^{i\lambda}|^2}.$$

*Then we have*

$$\mathrm{tr}\, \Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\dot{\Sigma}_T = -\frac{\partial^2 \log\det \Sigma_T}{\partial\theta^2} + \mathrm{tr}\, \Sigma_T^{-1}\ddot{\Sigma}_T, \tag{5.20}$$

$$\log\det \Sigma_T = 2\log(1-\alpha\beta) - \log(1-\alpha^2) - \log(1-\beta^2)$$
$$+ T\log\sigma^2 + O(\beta^{2T}), \tag{5.21}$$

*where $\dot{\Sigma}_T$ and $\ddot{\Sigma}_T$ are the $T\times T$-Toeplitz type matrices whose $(m,n)$th elements are given by $\int_{-\pi}^{\pi} e^{i(m-n)\lambda}(\partial/\partial\theta) f_\theta(\lambda)\, d\lambda$ and $\int_{-\pi}^{\pi} e^{i(m-n)\lambda}(\partial^2/\partial\theta^2) f_\theta(\lambda)\, d\lambda$, respectively.*

Put $\Sigma_T^{-1}=\{m_{rs}\}$, $r$, $s=1,\ldots, T$. Galbraith and Galbraith [11] gave the exact expressions of $m_{rs}$ for ARMA(1, 1). From their exact expressions we get

$$\begin{aligned} m_{rr} &= \frac{(\beta-\alpha)^2}{\sigma^2(1-\beta^2)}\left[1-\beta^{2(r-1)}-\beta^{2(T-r)}\right]+\frac{1}{\sigma^2}+O(\beta^T), \\ &= \tilde{m}_{rr}+O(\beta^T) \qquad \text{say } r=1,\ldots, T, \end{aligned} \tag{5.22}$$

$$\begin{aligned} m_{rs} &= \frac{\beta^{s-r-1}(\beta-\alpha)(1-\alpha\beta)}{\sigma^2(1-\beta^2)}-\frac{(\beta-\alpha)^2}{\sigma^2(1-\beta^2)}\beta^{r+s-2} \\ &\quad -\frac{(\beta-\alpha)^2}{\sigma^2(1-\beta)}\beta^{2T-s-r}+O(\beta^T), \\ &= \tilde{m}_{rs}+O(\beta^T) \qquad \text{say } 1\leqslant r<s\leqslant T. \end{aligned} \tag{5.23}$$

Now we can do the following evaluation:

$$\operatorname{tr} \Sigma_T^{-1}\ddot{\Sigma}_T=\sum_{r=1}^{T}\tilde{m}_{rr}a_{rr}+2\sum_{r=1}^{T-1}\sum_{s=r+1}^{T}\tilde{m}_{rs}a_{rs}+O(T^2\beta^T), \tag{5.24}$$

where $a_{rs}$ is the $(r,s)$th element of $\ddot{\Sigma}_T$. Using the residue theorem and Lemma 7 we have

PROPOSITION 1. *For the spectral density*

$$f_\theta(\lambda)=\frac{\sigma^2}{2\pi}\frac{|1-\beta e^{i\lambda}|^2}{|1-\alpha e^{i\lambda}|^2},$$

*we have*

$$\begin{aligned} I(\sigma^2)&=\frac{1}{2\sigma^4}, & I(\beta)&=\frac{1}{1-\beta^2}, & I(\alpha)&=\frac{1}{1-\alpha^2}, \\ J(\sigma^2)&=-\frac{1}{\sigma^6}, & J(\beta)&=\frac{4\beta}{(1-\beta^2)^2}, & J(\alpha)&=\frac{-2\alpha}{(1-\alpha^2)^2}, \\ K(\sigma^2)&=\frac{1}{\sigma^6}, & K(\beta)&=\frac{-6\beta}{(1-\beta^2)^2}, & K(\alpha)&=\frac{6\alpha}{(1-\alpha^2)^2}, \\ L(\sigma^2)&=\frac{3}{\sigma^8}, & L(\beta)&=\frac{6+18\beta^2}{(1-\beta^2)^3}, & L(\alpha)&=0, \end{aligned}$$

$$M(\sigma^2)=\frac{2}{\sigma^8},\qquad M(\beta)=\frac{6+14\beta^2}{(1-\beta^2)^3},\qquad M(\alpha)=\frac{2+2\alpha^2}{(1-\alpha^2)^3},$$

$$N(\sigma^2)=-\frac{2}{\sigma^8},\qquad N(\beta)=\frac{-8-20\beta^2}{(1-\beta^2)^3},\qquad N(\alpha)=\frac{-4-8\alpha^2}{(1-\alpha^2)^3},$$

$$H(\sigma^2)=\frac{3}{\sigma^8},\qquad H(\beta)=\frac{6(3+7\beta^2)}{(1-\beta^2)^3},\qquad H(\alpha)=\frac{6(3+7\alpha^2)}{(1-\alpha^2)^3},$$

$$\Delta(\sigma^2)=0,\qquad \Delta(\beta)=\frac{\alpha^2(1-5\beta^2)+\alpha(4\beta+4\beta^3)-3\beta^2-1}{(1-\beta^2)^2\,(1-\alpha\beta)^2},$$

$$\Delta(\alpha)=\frac{\beta^2(3\alpha^2-1)-4\alpha^3\beta+3\alpha^2-1}{(1-\alpha^2)^2\,(1-\alpha\beta)^2}.$$

Using the above results we get

THEOREM 2. *For the* ARMA *spectral density model*

$$f_\theta(\lambda)=\frac{\sigma^2}{2\pi}\frac{|1-\beta e^{i\lambda}|^2}{|1-\alpha e^{i\lambda}|^2},$$

*we have the following third-order bound distributions for $\theta=\sigma^2$, $\beta$, and $\alpha$:*

$$F^{(3)}_{\sigma^2}(y)=\Phi(y)-\phi(y)\left[\frac{\sqrt{2}y^2}{3\sqrt{T}}+\frac{5y}{18T}-\frac{7y^3}{18T}+\frac{y^5}{9T}\right], \tag{5.25}$$

$$F^{(3)}_{\beta}(y)=\Phi(y)-\phi(y)\left[\frac{\alpha^2(3\beta^4+13\beta^2-2)-2\alpha(7\beta^3+7\beta)+9\beta^2+5}{4T(1-\beta^2)(1-\alpha\beta)^2}\,y\right], \tag{5.26}$$

$$\begin{aligned}F^{(3)}_{\alpha}(y)=\Phi(y)-\phi(y)\Bigg[&-\frac{\alpha y^2}{\sqrt{T(1-\alpha^2)}}\\&+\frac{\{(3\alpha^4-3\alpha^2+2)\,\beta^2+(2\alpha^3-6\alpha)\,\beta+5-3\alpha^2\}}{4T(1-\alpha\beta)^2\,(1-\alpha^2)}\,y\\&-\frac{(1+2\alpha^2)\,y^3}{2T(1-\alpha^2)}+\frac{\alpha^2y^5}{2T(1-\alpha^2)}\Bigg].\end{aligned} \tag{5.27}$$

*Remark* 1. *In the special case of $\beta=0$, i.e., an autoregressive model of order* 1, *the above bound distribution $F^{(3)}_\alpha(y)$ becomes*

$$\Phi(y)+\phi(y)\left[\frac{\alpha y^2}{\sqrt{T(1-\alpha^2)}}+\frac{(3\alpha^2-5)\,y}{4T(1-\alpha^2)}+\frac{(1+2\alpha^2)\,y^3}{2T(1-\alpha^2)}-\frac{\alpha^2y^5}{2T(1-\alpha^2)}\right],$$

*which coincides with the result of Fujikoshi and Ochi* [10].

## 6. Third-Order Asymptotic Properties of MLE

In this section we shall investigate the third-order asymptotic properties of MLE. Here we shall get interesting results. That is, appropriately modified (to be third-order AMU) maximum likelihood estimators are not always third-order asymptotically efficient in $A_3$.

Now we use the following notations throughout this section:

$$Z^{(1)}(\theta)=\frac{1}{\sqrt{T}}\frac{\partial}{\partial\theta}G(\theta), \tag{6.1}$$

$$Z^{(2)}(\theta)=\frac{1}{\sqrt{T}}\left\{\frac{\partial^2}{\partial\theta^2}G(\theta)-E_\theta\frac{\partial^2}{\partial\theta^2}G(\theta)\right\}, \tag{6.2}$$

$$Z^{(3)}(\theta)=\frac{1}{\sqrt{T}}\left\{\frac{\partial^3}{\partial\theta^3}G(\theta)-E_\theta\frac{\partial^3}{\partial\theta^3}G(\theta)\right\}, \tag{6.3}$$

where $G(\theta)$ is defined by (5.1). For simplicity we sometimes use $Z^{(1)}$, $Z^{(2)}$, $Z^{(3)}$ instead of $Z^{(1)}(\theta)$, $Z^{(2)}(\theta)$, $Z^{(3)}(\theta)$, respectively. For a while we develop the discussion by using the argument similar to that of Bhattacharya and Ghosh [4] and Taniguchi [16]. Consider the equation

$$\begin{aligned}0=T^{-1}\frac{\partial}{\partial\theta}G(\theta_0)&+T^{-1}(\theta-\theta_0)\frac{\partial^2}{\partial\theta^2}G(\theta_0)\\&+(2T)^{-1}(\theta-\theta_0)^2\frac{\partial^3}{\partial\theta^3}G(\theta_0)\\&+(6T)^{-1}(\theta-\theta_0)^3\frac{\partial^4}{\partial\theta^4}G(\theta_0)+R_T(\theta),\end{aligned} \tag{6.4}$$

where $R_T(\theta)$ is the usual remainder in the Taylor expansion, for which it holds that

$$|R_T(\theta)|\leqslant\frac{1}{24T}|\theta-\theta_0|^4\sup_{|\theta'-\theta|\leqslant|\theta-\theta_0|}\left|\frac{\partial^5}{\partial\theta^5}G(\theta')\right|. \tag{6.5}$$

In view of (4.26) and (4.27) in Lemma 5, we can see that there exist positive constants $d_1$ and $d_2$ such that

$$P_{\theta_0}[Z^{(1)}(\theta_0)>d_1T^\alpha]=o(T^{-1}), \tag{6.6}$$

$$P_{\theta_0}[Z^{(2)}(\theta_0)>d_1T^\alpha]=o(T^{-1}), \tag{6.7}$$

$$P_{\theta_0}[Z^{(3)}(\theta_0)>d_1T^\alpha]=o(T^{-1}), \tag{6.8}$$

$$P_{\theta_0}\left[\left|\frac{1}{\sqrt{T}}\left\{\frac{\partial^4}{\partial\theta^4}G(\theta_0)-E_{\theta_0}\frac{\partial^4}{\partial\theta^4}G(\theta_0)\right\}\right|>d_1T^{\alpha}\right]=o(T^{-1}), \tag{6.9}$$

$$P_{\theta_0}[|R_T(\theta)|>|\theta-\theta_0|^4\{d_2+d_1T^{-1/2+\alpha}\}]=o(T^{-1}), \tag{6.10}$$

for any $\alpha$ $(0<\alpha<\frac{1}{2})$. Therefore, on a set having $P_{\theta_0}$-probability at least $1-o(T^{-1})$, for some constants $d_3$ and $d_4>0$ we can rewrite (6.4) as

$$\begin{aligned}\theta-\theta_0=(I(\theta_0)+\eta_T)^{-1}\Big[\delta_T+(2T)^{-1}(\theta-\theta_0)^2\frac{\partial^3}{\partial\theta^3}G(\theta_0)\\ +\frac{1}{6T}(\theta-\theta_0)^3\frac{\partial^4}{\partial\theta^4}G(\theta_0)+d_3|\theta-\theta_0|^4\zeta_T\Big],\end{aligned} \tag{6.11}$$

where $\eta_T$ and $\delta_T$ are random variables whose absolute values are less than $d_4T^{-1/2+\alpha}$ and $\zeta_T$ is a random variable whose absolute value is less than one. There exist a sufficiently large $d_5>0$ and an integer $T_0$ such that if $T>T_0$ and $|\theta-\theta_0|\leqslant d_5T^{-1/2+\alpha}$, the right-hand side of (6.11) is less than $d_5T^{-1/2+\alpha}$. Applying the Brouwer fixed point theorem to the right-hand side of (6.11) we have the following proposition.

PROPOSITION 2. *There exists a statistic $\hat{\theta}_{\mathrm{ML}}$ such that*

$$\begin{aligned}P_{\theta_0}[|\hat{\theta}_{\mathrm{ML}}-\theta_0|<d_5T^{-1/2+\alpha},\ \hat{\theta}_{\mathrm{ML}}\text{ solves (6.4)}]\\ =1-o(T^{-1}).\end{aligned} \tag{6.12}$$

Putting $V_T=\sqrt{T}(\hat{\theta}_{\mathrm{ML}}-\theta_0)$ we have

$$\begin{aligned}0=Z^{(1)}(\theta_0)+T^{-1/2}Z^{(2)}(\theta_0)V_T-I_T(\theta_0)V_T+2^{-1}T^{-3/2}\left\{\frac{\partial^3}{\partial\theta^3}G(\theta_0)\right\}V_T^2\\ +\frac{1}{6T^2}\left\{\frac{\partial^4}{\partial\theta^4}G(\theta_0)\right\}V_T^3+\frac{1}{24T^{5/2}}\left\{\frac{\partial^5}{\partial\theta^5}G(\theta^*)\right\}V_T^4,\end{aligned} \tag{6.13}$$

where $|\theta^*-\theta_0|\leqslant|\hat{\theta}_{\mathrm{ML}}-\theta_0|$. We rewrite (6.13) as

$$\begin{aligned}V_T=\frac{Z^{(1)}(\theta_0)}{I_T(\theta_0)}+\frac{Z^{(2)}(\theta_0)V_T}{\sqrt{T}I_T(\theta_0)}+\frac{1}{2I_T(\theta_0)\sqrt{T}}\left\{\frac{1}{T}\frac{\partial^3}{\partial\theta^3}G(\theta_0)\right\}V_T^2\\ +\frac{1}{6TI_T(\theta_0)}\left\{\frac{1}{T}\frac{\partial^4}{\partial\theta^4}G(\theta_0)\right\}V_T^3+\frac{1}{24I_T(\theta_0)T^{3/2}}\left\{\frac{1}{T}\frac{\partial^5}{\partial\theta^5}G(\theta^*)\right\}V_T^4.\end{aligned} \tag{6.14}$$

Notice that

$$E_\theta T^{-1}\frac{\partial^3}{\partial\theta^3}G(\theta)=-3J(\theta)-K(\theta)+O(T^{-1}),\tag{6.15}$$

$$E_\theta T^{-1}\frac{\partial^4}{\partial\theta^4}G(\theta)=-4L(\theta)-3M(\theta)-6N(\theta)-H(\theta)+O(T^{-1}).\tag{6.16}$$

Substituting

$$\begin{aligned}U_T(\theta_0)=&\frac{Z^{(1)}}{I_T}+\frac{1}{\sqrt{T}I^2}\left\{Z^{(1)}Z^{(2)}-\frac{3J+K}{2I}Z^{(1)^2}\right\}\\&+\frac{1}{TI^3}\left\{Z^{(1)}Z^{(2)^2}+\frac{1}{2}Z^{(1)^2}Z^{(3)}-\frac{3(3J+K)}{2I}Z^{(1)^2}Z^{(2)}\right.\\&\left.+\frac{(3J+K)^2}{2I^2}Z^{(1)^3}-\frac{4L+3M+6N+H}{6I}Z^{(1)^3}\right\}\end{aligned}$$

for $V_T$ in the right-hand side of (6.14), and using (6.6)-(6.10), (6.12), (6.15), (6.16), we have the following proposition:

PROPOSITION 3. *We have the following stochastic expansion*

$$\sqrt{T}(\hat{\theta}_{\mathrm{ML}}-\theta_0)=U_T(\theta_0)+T^{-3/2}\xi_T,$$

*where $\xi_T$ satisfies $P_{\theta_0}\{|\xi_T|>\rho_T\sqrt{T}\}=o(T^{-1})$ for some sequence $\rho_T\to 0$, $\rho_T\sqrt{T}\to\infty$ as $T\to\infty$.*

*Remark 2. By Lemma 6, the Edgeworth expansion for $\sqrt{T}(\hat{\theta}_{\mathrm{ML}}-\theta_0)$ (up to order $T^{-1}$) is equal to that for $U_T(\theta_0)$. Thus we have only to derive the Edgeworth expansion for $U_T(\theta_0)$.*

It is not difficult to show

$$EU_T=-\frac{J+K}{2\sqrt{T}I^2}+o(T^{-1}),\tag{6.17}$$

$$\operatorname{Var}U_T=I^{-1}-\frac{\Delta}{I^2T}+\frac{7J^2+14JK+5K^2}{2I^4T}-\frac{L+4N+H}{I^3T}+o(T^{-1}),\tag{6.18}$$

$$\operatorname{cum}(U_T,U_T,U_T)=-\frac{3J+2K}{I^3\sqrt{T}}+o(T^{-1}),\tag{6.19}$$

$$\operatorname{cum}(U_T,U_T,U_T,U_T)=\frac{12(2J+K)(J+K)}{I^5T}-\frac{4L+12N+3H}{I^4T}+o(T^{-1}),\tag{6.20}$$

$$\operatorname{cum}^{(J)}(U_T,\dots,U_T)=O(T^{-J/2+1})\qquad\text{for } J\geqslant 5.\tag{6.21}$$

Thus we have the following theorem:

THEOREM 3.

$$P_\theta^T\{\sqrt{T}I(\hat{\theta}_{\mathrm{ML}}-\theta)\leqslant x\}=\Phi(x)-\phi(x)\Big\{\frac{\alpha_1}{\sqrt{T}}+\frac{\gamma_1}{6\sqrt{T}}(x^2-1)$$

$$+\frac{1}{2}\left(\frac{\rho_2}{T}+\frac{\alpha_1^2}{T}\right)x+\left(\frac{\delta_1}{24T}+\frac{\alpha_1\gamma_1}{6T}\right)(x^3-3x)$$

$$+\frac{\gamma_1^2}{72T}(x^5-10x^3+15x)]+o(T^{-1}), \tag{6.22}$$

*where*

$$\alpha_1=-\frac{J+K}{2I^{3/2}},\qquad \rho_2=\frac{7J^2+14JK+5K^2}{2I^3}-\frac{L+4N+H}{I^2}-\frac{\Delta}{I},$$

$$\gamma_1=-\frac{3J+2K}{I^{3/2}},\qquad \delta_1=\frac{12(2J+K)(J+K)}{I^3}-\frac{4L+12N+3H}{I^2}.$$

*In the special case of*

$$f_\theta(\lambda)=\frac{\sigma^2}{2\pi}\frac{|1-\beta e^{i\lambda}|^2}{|1-\alpha e^{i\lambda}|^2},$$

*we have*

$$P_\theta^T\left\{\sqrt{\frac{T}{2\sigma^4}}(\hat{\sigma}_{\mathrm{ML}}^2-\sigma^2)\leqslant x\right\}=\Phi(x)-\phi(x)\Big\{\frac{\sqrt{2}}{3\sqrt{T}}(x^2-1)+\frac{x}{6T}$$

$$-\frac{11x^3}{18T}+\frac{x^5}{9T}\Big\}+o(T^{-1}), \tag{6.23}$$

$$P_\theta^T\left\{\sqrt{\frac{T}{1-\beta^2}}(\hat{\beta}_{\mathrm{ML}}-\beta)\leqslant x\right\}$$

$$=\Phi(x)-\phi(x)\Big\{\frac{\beta}{\sqrt{T}\sqrt{1-\beta^2}}$$

$$+\frac{\alpha^2(\beta^4+17\beta^2-2)-\alpha(22\beta+10\beta^3)+7\beta^2+9}{4T(1-\beta^2)(1-\alpha\beta)^2}x$$

$$+\frac{3-\beta^2}{4T(1-\beta^2)}x^3\Big\}+o(T^{-1}), \tag{6.24}$$

$$P_\theta^T\left\{\sqrt{\frac{T}{1-\alpha^2}}(\hat{\alpha}_{\mathrm{ML}}-\alpha)\leqslant x\right\}=\Phi(x)-\phi(x)\left\{\frac{-\alpha(x^2+1)}{\sqrt{T}\sqrt{1-\alpha^2}}\right.$$
$$+\frac{\beta^2(\alpha^4-7\alpha^2+2)+\beta(6\alpha^3+2\alpha)+1-5\alpha^2}{4T(1-\alpha^2)(1-\alpha\beta)^2}x$$
$$\left.-\frac{\alpha^2+1}{4T(1-\alpha^2)}x^3+\frac{\alpha^2x^5}{2T(1-\alpha^2)}\right\}+o(T^{-1}),\tag{6.25}$$

*where $\hat{\sigma}^2_{\mathrm{ML}}$, $\hat{\beta}_{\mathrm{ML}}$, and $\hat{\alpha}_{\mathrm{ML}}$ are the maximum likelihood estimators of $\sigma^2$, $\beta$, and $\alpha$, respectively.*

*Remark* 3. *In the special case of $\beta=0$, i.e., an autoregressive model of order* 1, *the right-hand side of* (6.25) *becomes*

$$\Phi(x)-\phi(x)\left[-\frac{\alpha(x^2+1)}{\sqrt{T}\sqrt{1-\alpha^2}}+\frac{1-5\alpha^2}{4T(1-\alpha^2)}x-\frac{\alpha^2+1}{4T(1-\alpha^2)}x^3\right.$$
$$\left.+\frac{\alpha^2x^5}{2T(1-\alpha^2)}\right]+o(T^{-1}),\tag{6.26}$$

*which coincides with the result of Fujikoshi and Ochi* [10].

Also we can evaluate he mean square errors of estimators up to $T^{-1}$-order.

THEOREM 4.

$$E\{\sqrt{TI}(\hat{\theta}_{\mathrm{ML}}-\theta)\}^2$$
$$=1+\frac{1}{T}\left[\frac{15J^2+30JK+11K^2}{4I^3}-\frac{L+4N+H}{I^2}-\frac{\Delta}{I}\right]+o(T^{-1}).\tag{6.27}$$

*In the special case of*

$$f_\theta(\lambda)=\frac{\sigma^2}{2\pi}\frac{|1-\beta e^{i\lambda}|^2}{|1-\alpha e^{i\lambda}|^2},$$

*we have*

$$E\left\{\sqrt{\frac{T}{2\sigma^4}}(\hat{\sigma}^2_{\mathrm{ML}}-\sigma^2)\right\}^2=1+o(T^{-1}),\tag{6.28}$$

$$E\left\{\sqrt{\frac{T}{1-\beta^2}}(\hat\beta_{\mathrm{ML}}-\beta)\right\}^2$$
$$=1+\frac{1}{T}\left\{\frac{\alpha^2(-\beta^4+13\beta^2-1)+\alpha(-2\beta^3-20\beta)+2\beta^2+9}{(1-\beta^2)(1-\alpha\beta)^2}\right\}+o(T^{-1}) \tag{6.29}$$

$$E\left\{\sqrt{\frac{T}{1-\alpha^2}}(\hat\alpha_{\mathrm{ML}}-\alpha)\right\}^2$$
$$=1+\frac{1}{T}\left\{\frac{\beta^2(14\alpha^4-5\alpha^2+1)+\beta(4\alpha-24\alpha^3)+11\alpha^2-1}{(1-\alpha^2)(1-\alpha\beta)^2}\right\}+o(T^{-1}). \tag{6.30}$$

For i.i.d. multinomial case the evaluation of the type (6.27) for the variance of the maximum likelihood estimator has been studied earlier by Rao (1962) who introduced the concept of the second order efficiency which corresponds to our third order efficiency in a class $D$ which will be defined in the next section.

Now we shall discuss the third-order asymptotic efficiency of the maximum likelihood estimators. To do so we modify $\hat\theta_{\mathrm{ML}}$ to be third order AMU. That is, we put

$$\hat\theta^*_{\mathrm{ML}}=\hat\theta_{\mathrm{ML}}+\frac{K(\hat\theta_{\mathrm{ML}})}{6TI^2(\hat\theta_{\mathrm{ML}})}. \tag{6.31}$$

Noting that

$$\sqrt{T}(\hat\theta^*_{\mathrm{ML}}-\theta)=\sqrt{T}(\hat\theta_{\mathrm{ML}}-\theta)+\frac{K}{6\sqrt{T}I^2}+\frac{\sqrt{T}(\hat\theta_{\mathrm{ML}}-\theta)}{6T}\frac{\partial}{\partial\theta}\left\{\frac{K}{I^2}\right\}+o_p(T^{-1}),$$

we can show that

$$T\,\mathrm{Var}(\hat\theta^*_{\mathrm{ML}})=\left\{1+\frac{(3N+H)I-2K(2J+K)}{3TI^3}\right\}T\,\mathrm{Var}(\hat\theta_{\mathrm{ML}})+o(T^{-1}). \tag{6.32}$$

Putting $U^*_T=\sqrt{TI}(\hat\theta^*_{\mathrm{ML}}-\theta)$, we have

$$EU^*_T=-\frac{3J+2K}{6I^{3/2}\sqrt{T}}+o(T^{-1}), \tag{6.33}$$

$$\mathrm{Var}\,U^*_T=1-\frac{\Delta}{IT}+\frac{-3L-9N-2H}{3I^2T}+\frac{21J^2+34JK+11K^2}{6I^3T}+o(T^{-1}), \tag{6.34}$$

$$\operatorname{cum}(U_T^*, U_T^*, U_T^*) = -\frac{3J+2K}{I^{3/2}\sqrt{T}} + o(T^{-1}), \tag{6.35}$$

$$\operatorname{cum}(U_T^*, U_T^*, U_T^*, U_T^*) = \frac{12(2J+K)(J+K)}{I^3 T} - \frac{4L+12N+3H}{I^2 T} + o(T^{-1}), \tag{6.36}$$

$$\operatorname{cum}^{(J)}(U_T^*, \dots, U_T^*) = O(T^{-J/2+1}) \qquad \text{for } J \geqslant 5. \tag{6.37}$$

Thus we have

THEOREM 5.

$$\begin{aligned} &P_\theta^T\left\{\sqrt{TI}(\hat{\theta}_{\mathrm{ML}}^* - \theta) \leqslant y\right\} \\ &= \Phi(y) - \phi(y)\left[-\frac{3J+2K}{6\sqrt{T}I^{3/2}}y^2\right. \\ &\quad + \frac{1}{T}\left\{-\frac{\Delta}{2I} - \frac{K^2}{36I^3} + \frac{H}{24I^2}\right\}y + \frac{1}{T}\left\{\frac{3KJ+K^2}{18I^3} - \frac{4L+12N+3H}{24I^2}\right\}y^3 \\ &\quad \left. + \frac{(3J+2K)^2}{72TI^3}y^5\right] + o(T^{-1}). \end{aligned} \tag{6.38}$$

*Thus the difference between the third-order bound distribution $F_\theta^{(3)}(y)$ given by* (5.18) *and* (6.38) *is*

$$\begin{aligned} &F_\theta^{(3)}(y) - P_\theta^T\{\sqrt{TI}\,(\hat{\theta}_{\mathrm{ML}}^* - \theta) \leqslant y\} \\ &= \frac{y^3\phi(y)}{8TI^3}(MI - J^2) + o(T^{-1}) \qquad \textit{for } y > 0. \end{aligned} \tag{6.39}$$

*In the special case of*

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi}\frac{|1-\beta e^{i\lambda}|^2}{|1-\alpha e^{i\lambda}|^2} \qquad \textit{for } \hat{\sigma}_{\mathrm{ML}}^{2*} = \left(1 + \frac{2}{3T}\right)\hat{\sigma}_{\mathrm{ML}}^2,$$

$\hat{\beta}_{\mathrm{ML}}^* = (1 - 1/T)\,\hat{\beta}_{\mathrm{ML}}$ *and* $\hat{\alpha}_{\mathrm{ML}}^* = (1 + 1/T)\,\hat{\alpha}_{\mathrm{ML}}$ *which are third-order AMU, we have*

$$F_{\sigma^2}^{(3)}(y) - P_{\sigma^2}^T\left\{\sqrt{\frac{T}{2\sigma^4}}\,(\hat{\sigma}_{\mathrm{ML}}^{2*} - \sigma^2) \leqslant y\right\} = o(T^{-1}), \tag{6.40}$$

(i.e., $\hat{\sigma}_{\mathrm{ML}}^{2*}$ *is third-order asymptotically efficient*),

$$\begin{aligned} &F_\beta^{(3)}(y) - P_\beta^T\left\{\sqrt{\frac{T}{1-\beta^2}}\,(\hat{\beta}_{\mathrm{ML}}^* - \beta) \leqslant y\right\} \\ &= \frac{3-\beta^2}{4(1-\beta^2)\,T}\phi(y)\,y^3 > 0 \qquad \textit{for } y > 0, \end{aligned} \tag{6.41}$$

(i.e., $\hat{\beta}^*_{\text{ML}}$ *is not third-order asymptotically efficient*),

$$F_\alpha^{(3)}(y) - P_\alpha^T\left\{\sqrt{\frac{T}{1-\alpha^2}}\,(\hat{\alpha}^*_{\text{ML}} - \alpha) \leqslant y\right\} = \frac{y^3\phi(y)}{4T} > 0 \qquad \textit{for } y > 0, \tag{6.42}$$

(*i.e.*, $\hat{\alpha}^*_{\text{ML}}$ *is not third-order asymptotically efficient*).

Also we have the following theorem:

THEOREM 6. *The modified maximum likelihood estimator* $\hat{\theta}^*_{\text{ML}} = \hat{\theta}_{\text{ML}} + K(\hat{\theta}_{\text{ML}})/6TI^2(\hat{\theta}_{\text{ML}})$ *is third-order asymptotically efficient if and only if the spectral density* $f_\theta(\lambda)$ *satisfies the following differential equation*;

$$\frac{\partial^2 \log f_\theta(\lambda)}{\partial\theta^2} - \left(\frac{\partial \log f_\theta(\lambda)}{\partial\theta}\right)^2 + c(\theta)\left(\frac{\partial \log f_\theta(\lambda)}{\partial\theta}\right) = 0, \tag{6.43}$$

*where* $c(\theta)$ *is a function which depends only on* $\theta$. *The condition* (6.43) *is equivalent that the spectral density* $f_\theta(\lambda)$ *is parameterized as*

$$f_\theta(\lambda) = s(\lambda)\exp\left[\int \frac{\exp\left(\int -c(\theta)\,d\theta\right)}{-\int \exp\left(\int -c(\theta)\,d\theta\right)d\theta + b(\lambda)}\,d\theta\right], \tag{6.44}$$

*where* $s(\cdot)$, $b(\cdot)$ *are functions which depend only on* $\lambda$.

*Proof.* By (6.39) we can see that $\hat{\theta}^*_{\text{ML}}$ is third-order asymptotically efficient if and only if $MI - J^2 = 0$. By Schwarz's inequality we have

$$\begin{aligned} MI &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\left\{2\left(\frac{\partial f_\theta}{\partial\theta}\right)^2 f_\theta^{-2} - \frac{\partial^2 f_\theta}{\partial\theta^2} f_\theta^{-1}\right\}^2 d\lambda \times \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\partial f_\theta}{\partial\theta}\right)^2 f_\theta^{-2}\,d\lambda \\ &\geqslant \left[\frac{1}{4\pi}\int_{-\pi}^{\pi}\left\{\frac{\partial^2 f_\theta}{\partial\theta^2} f_\theta^{-1} - 2\left(\frac{\partial f_\theta}{\partial\theta}\right)^2 f_\theta^{-2}\right\}\left\{\frac{\partial f_\theta}{\partial\theta} f_\theta^{-1}\right\} d\lambda\right]^2 = J^2, \end{aligned} \tag{6.45}$$

where the equality holds if and only if

$$\frac{\partial^2 f_\theta}{\partial\theta^2} f_\theta^{-1} - 2\left(\frac{\partial f_\theta}{\partial\theta}\right)^2 f_\theta^{-2} = -c(\theta)\frac{\partial f_\theta}{\partial\theta} f_\theta^{-1}, \tag{6.46}$$

where $c(\theta)$ depends only on $\theta$. The above (6.46) implies (6.43). From (6.43) we have

$$\frac{\partial \log f_\theta(\lambda)}{\partial\theta} = \left\{\exp\int -c(\theta)\,d\theta\right\}\left\{\int -e^{\int -c(\theta)d\theta}\,d\theta + b(\lambda)\right\}^{-1},$$

which implies (6.44). ∎

## 7. Third-Order Asymptotic Efficiency of MLE in $D$

As we saw in the previous sections the maximum likelihood estimator is not always third-order asymptotically efficient in the class $A_3$. However, in this section we shall get a unified result. That is, if we confine ourselves to a class of estimators $D \subset A_3$, then we can show that the maximum likelihood estimator which is modified to be third-order AMU, is third-order asymptotically efficient in $D$. Hereafter we use the notations in Section 4. We shall state the results without proofs because we extend the results of Taniguchi [17] to the case when unknown parameter $\boldsymbol{\theta}$ is a vector.

Let $\mathbf{X}_T = (X_1, ..., X_T)'$ be a stretch from the series $\{X_t\}$ which satisfies Assumptions 1–4. We set

$$\mathbf{Z}^{(1)} = (Z_1, ..., Z_p)', \qquad \mathbf{Z}^{(2)} = (Z_{ij}), \qquad \mathbf{U} = I(\boldsymbol{\theta})^{-1} \mathbf{Z}^{(1)},$$

and

$$E_{\boldsymbol{\theta}} \mathbf{Z}^{(1)} \mathbf{Z}^{(1)'} = I(\boldsymbol{\theta}) + \frac{\Delta(\boldsymbol{\theta})}{T} + o(T^{-1}).$$

Let $S$ be the class of the estimators $\hat{\boldsymbol{\theta}}_T$ which are asymptotically expanded as

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) = \mathbf{U} + \mathbf{Q}/\sqrt{T} + o_p(T^{-1/2}), \tag{7.1}$$

where $\mathbf{Q} = (Q_1, ..., Q_p)' = O_p(1)$. We assume that $\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})$ has the Edgeworth expansion up to the order $T^{-1}$ and that

$$E_{\boldsymbol{\theta}} \sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) = \boldsymbol{\mu}/\sqrt{T} + o(T^{-1}), \tag{7.2}$$

where $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)' = E_{\boldsymbol{\theta}} \mathbf{Q}$. Putting

$$\mathbf{S} = (S_1, ..., S_p)' = \sqrt{T}(\hat{\boldsymbol{\theta}}_T - E_{\boldsymbol{\theta}} \hat{\boldsymbol{\theta}}_T),$$

we can show that

$$E_{\boldsymbol{\theta}} S_i S_j = I^{ij} - \frac{\eta_{ij}}{T} + \frac{1}{T} D_i \mu_j + \frac{1}{T} D_j \mu_i + \frac{1}{T} \operatorname{Cov}(Q_i, Q_j) + o(T^{-1}), \tag{7.3}$$

where $I^{ij}$ and $\eta_{ij}$ are the $(i, j)$th elements of $I(\boldsymbol{\theta})^{-1}$ and $I(\boldsymbol{\theta})^{-1} \Delta(\boldsymbol{\theta}) I(\boldsymbol{\theta})^{-1}$, respectively, and

$$D_i = \sum_{k=1}^{p} I^{ik} \frac{\partial}{\partial \theta_k} \qquad \text{(differential operator)}.$$

We denote

$$\beta_{ijk} = -\sum_{i',j',k'=1}^{p} I^{ii'}I^{jj'}I^{kk'}\{2K_{i'j'k'} + J_{i'j'k'} + J_{j'k'i'} + J_{k'i'j'}\},$$

$$A_{ijk} = E_{\boldsymbol{\theta}} U_i \tilde{Q}_j \tilde{Q}_k + E_{\boldsymbol{\theta}} U_j \tilde{Q}_i \tilde{Q}_k + E_{\boldsymbol{\theta}} U_k \tilde{Q}_i \tilde{Q}_j,$$

where $\tilde{Q}_i = Q_i - \mu_i$. Then we can show that

$$E_{\boldsymbol{\theta}} S_i S_j S_k = \frac{1}{\sqrt{T}} \beta_{ijk} + \frac{1}{2T} A_{ijk} + o(T^{-1}). \tag{7.4}$$

We define

$$\begin{aligned}
\beta_{ijkm} = {} & -3H^{ijkm} - 2(N^{ijkm} + N^{ikjm} + N^{imjk} + N^{jkim} \\
& + N^{jmik} + N^{kmij}) - (L^{ijkm} + L^{ijmk} + L^{ikmj} \\
& + L^{jkmi}) + \tfrac{1}{2}(\Gamma^{ijkm} + \Gamma^{jikm} + \Gamma^{ikjm} \\
& + \Gamma^{kijm} + \Gamma^{jmik} + \Gamma^{mjik} + \Gamma^{kmij} + \Gamma^{mkij} \\
& + \Gamma^{jkim} + \Gamma^{kjim} + \Gamma^{imjk} + \Gamma^{mijk}),
\end{aligned}$$

where

$$\begin{aligned}
\Gamma^{ijkm} &= \Sigma\Sigma\Sigma\Sigma\Sigma\Sigma^{p}_{i',j',k',m',n',n=1} I^{ii'}I^{jj'}I^{kk'}I^{mm'}I^{nn'} \\
&\quad \times (K_{i'j'n'} + J_{n'i'j'} + J_{k'i'n'})(2K_{k'm'n'} + J_{m'k'n'} + J_{k'i'n'}), \\
H^{ijkm} &= \Sigma\Sigma\Sigma\Sigma^{p}_{i',j',k',m'=1} I^{ii'}I^{jj'}I^{kk'}I^{mm'}H_{i'j'k'm'}, \\
N^{ijkm} &= \Sigma\Sigma\Sigma\Sigma^{p}_{i',j',k',m'=1} I^{ii'}I^{jj'}I^{kk'}I^{mm'}N_{i'j'k'm'},
\end{aligned}$$

and

$$L^{ijkm} = \Sigma\Sigma\Sigma\Sigma^{p}_{i',j',k',m'=1} I^{ii'}I^{jj'}I^{kk'}I^{mm'}L_{i'j'k'm'}.$$

Then we can show

$$\operatorname{cum}(S_i, S_j, S_k, S_m) = \frac{1}{T}\beta_{ijkm} + o(T^{-1}). \tag{7.5}$$

Remembering (3.7), and noting (7.2)–(7.5) we get the following asymptotic expansion:

$$\begin{aligned}
& P^T_{\boldsymbol{\theta}}\{\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \in C\} \\
&= \int \cdots \int_C N(y; I(\boldsymbol{\theta})^{-1}) \Big[ 1 + \Sigma_i \frac{\mu_i}{\sqrt{T}} H_i(y) \\
&\quad + \frac{1}{2T}\Sigma_{i,j}\{D_i\mu_j + D_j\mu_i - \eta_{ij} + \operatorname{Cov}(Q_i, Q_j) + \mu_i\mu_j\}\, H_{ij}(y)
\end{aligned}$$

$$+\Sigma_{i,j,k}\left\{\frac{\beta_{ijk}}{6\sqrt{T}}+\frac{A_{ijk}}{12T}\right\}H_{ijk}(y)$$

$$+\Sigma_{i,j,k,m}\left\{\frac{\beta_{ijkm}}{24T}+\frac{\mu_i\beta_{jkm}}{6T}\right\}H_{ijkm}(y)$$

$$\left.+\frac{1}{72T}\Sigma_{i,j,k,i',j',k'}\beta_{ijk}\beta_{i'j'k'}H_{ijki'j'k'}(y)\right]dy+o(T^{-1}), \qquad (7.6)$$

where $C$ is a convex set in $R^p$.

Now we introduce a class $D$ ($\subset S$) of estimators which satisfy

$$A_{ijk}=o(1) \qquad \text{for } i,j,k=1,\dots,p. \qquad (7.7)$$

As the following two propositions will show, this class $D$ is a natural one.

PROPOSITION 4. *The maximum likelihood estimator $\hat{\boldsymbol{\theta}}_{\mathrm{ML}}$ of $\boldsymbol{\theta}$ belongs to $D$.*

Let $\hat{\boldsymbol{\theta}}_{q\mathrm{ML}}$ be a quasi-maximum likelihood estimator, which maximizes the quasi-likelihood;

$$-\tfrac{1}{2}\Sigma_{j=0}^{T-1}\{\log f_{\boldsymbol{\theta}}(\lambda_j)+\P_T(\lambda_j)/f_{\boldsymbol{\theta}}(\lambda_j)\},$$

with respect to $\boldsymbol{\theta}$, where $\lambda_j=2\pi j/T$, and $\P_T(\lambda_j)=(1/2\pi T)|\Sigma_{t=1}^{T}X_te^{-i\lambda_jt}|^2$. Then we have

PROPOSITION 5. *The quasi-maximum likelihood estimator $\hat{\boldsymbol{\theta}}_{q\mathrm{ML}}$ belongs to $D$.*

In view of (7.6), if we modify $\hat{\boldsymbol{\theta}}_T\in D$ to be coordinate-wise third-order AMU, then $\boldsymbol{\mu}=(\mu_1,\dots,\mu_p)'$ is specified by $\beta_{ijk}$. Thus undetermined term for $\hat{\boldsymbol{\theta}}_T\in D\cap A_3$ (in (7.6)) is only $\mathrm{Cov}(Q_i,Q_j)$. If we get an estimator in $D\cap A_3$ which minimizes the matrix $\{\mathrm{Cov}(Q_i,Q_j)\}$, then it maximizes the concentration probability. Here we have

PROPOSITION 6. *Suppose that an estimator $\tilde{\boldsymbol{\theta}}_T$ belongs to D. The matrix $\{\mathrm{Cov}(Q_i,Q_j),\ i,\ j=1,\dots,p\}$ is minimized if $\sqrt{T}(\tilde{\boldsymbol{\theta}}_T-\boldsymbol{\theta})$ has the following stochastic expansion*:

$$\sqrt{T}(\tilde{\boldsymbol{\theta}}_T-\boldsymbol{\theta})=\mathbf{U}+\frac{1}{\sqrt{T}}I(\boldsymbol{\theta})^{-1}\mathbf{Z}^{(2)}\mathbf{U}$$

$$+\frac{1}{2\sqrt{T}}I(\boldsymbol{\theta})^{-1}R\ldots\circ\mathbf{U}\circ\mathbf{U}+\frac{1}{\sqrt{T}}\xi+o_p(T^{-1/2}), \qquad (7.8)$$

*where* $\xi$ *is a constant vector and* $R_{\ldots} = \{R_{ijk}\}$, $R_{ijk} = -K_{ijk} - J_{ijk} - J_{jki} - J_{kij}$, *and* $R_{\ldots} \circ U \circ U$ *is a p-dimensional column vector with* $i$th *component* $\Sigma_{j,k} R_{ijk} U_j U_k$. ∎

The stochastic expansion (7.8) is nothing but that of maximum likelihood estimator (see Proposition 3 for scalar case). Finally we get the followng theorem:

THEOREM 7. *If we modify the maximum likelihood estimator of* $\boldsymbol{\theta}$ *to be third-order AMU, then it is third-order asymptotically efficient in the class D.*


## ACKNOWLEDGMENTS

Some parts of this work were done during the author's stay in The Australian National University. The author is grateful to Professor E. J. Hannan for his kindness. Thanks are also extended to Professors C. R. Rao, P. R. Krishnaiah, K. Takeuchi, Y. Fujikoshi, and M. Akahira for thier encouragement and comments.


## REFERENCES


[1] AKAHIRA, M. (1975). A note on the second-order asymptotic efficiency of estimators in an autoregressive process. *Rep. Univ. Electro-Commun.* **26** 143–149.

[2] AKAHIRA, M. (1979). On the second-order asymptotic optimality of estimators in an autoregressive process. *Rep. Univ. Electro-Commun* **29** 213–218.

[3] AKAHIRA, M., AND TAKEUCHI, K. (1981). *Asymptotic Efficiency of Statistical Estimators; Concepts and Higher Order Asymptotic Efficiency.* Springer Lecture Notes in Statistics Vol. 7. Springer-Verlag, New York.

[4] BHATTACHARYA, R. N., AND GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

[5] BRILLINGER, D. R. (1969). Asymptotic properties of spectral estimates of second order. *Biometrika* **56** 375–390.

[6] BRILLINGER, D. R. (1975). *Time Series; Data Analysis and Theory.* Holt, New York.

[7] CHIBISOV, D. M. (1972). An asymptotic expansion for the distribution of a statistic admitting an asymptotic expansion. *Theory Probab. Its Appl.* **17** 620–630.

[8] CHIBISOV, D. M. (1973). An asymptotic expansion for a class of estimators containing maximum likelihood estimators. *Theory Probab. Its Appl.* **18** 295–303.

[9] FUJIKOSHI, Y. (1980). Asymptotic expansions for the distributions of the sample roots under nonnormality. *Biometrika* **67** 45–51.

[10] FUJIKOSHI, Y., AND OCHI, Y. (1984). Asymptotic properties of the maximum likelihood estimate in the first order autoregressive process. *Ann. Inst. Statist. Math.*, in press.

[11] GALBRAITH, R. F., AND GALBRAITH, J. I. (1974). On the inverses of some patterned matrices arising in the theory of stationary time series. *J. Appl. Probab.* **11** 63–71.

[12] HANNAN, E. J. (1970). *Multiple Time Series.* Wiley, New York.

[13] HAYAKAWA, T. (1975). The likelihood ratio criterion for a composite hypothesis under a local alternative. *Biometrika* **62** 451–460.

[14] MAGNUS, J. R., AND NEUDECKER, H. (1979). The commutation matrix; Some properties and applications. *Ann. Statist.* **7** 381–394.

[15] TANIGUCHI, M. (1983). On the second order asymptotic efficiency of estimators of Gaussian ARMA processes. *Ann. Statist.* **11** 157–169.

[16] TANIGUCHI, M. (1984). Validity of Edgeworth espansions for statistics of time series. *J. Time Series Anal.* **5** 37–51.

[17] TANIGUCHI, M. (1985). Third order efficiency of the maximum likelihood estimator in Gaussian autoregressive moving average processes. In *Statistical Theory and Data Analysis* (K. Marusita, Ed.), pp. 725–743. North-Holland, Amsterdam.

[18] WHITTLE, P. (1951). *Hypothesis Testing in Time Series Analysis.* Uppsala.

[19] FANG C., AND KRISHNAIAH, P. R. (1982). Asymptotic distributions of functions of the eigenvalues of some random matrices for non-normal populations. *J. Multivariate Anal.* **12** 36–63.

[20] RAO, C. R. (1962). Efficient estimates and optimum inference procedures in large samples. *J. R. Statist. Soc., Ser. B* **24** 46–72.